



Susman Godfrey L.L.P.  
a registered limited liability partnership



October 15, 2024

**VIA ECF**

Hon. Ona T. Wang  
Daniel Patrick Moynihan  
United States Courthouse  
500 Pearl St.  
New York, NY 10007

RE: *Authors Guild v. OpenAI Inc.*, 23-cv-8292 (S.D.N.Y.) and *Alter v. OpenAI Inc.*, 23-cv-10211 (S.D.N.Y.): Entry of a Deposition Coordination Protocol

Dear Judge Wang:

Pursuant to Rule II(b) of Your Honor's Individual Practices, Plaintiffs seek a conference to request the entry of their proposed deposition coordination protocol, attached hereto as Exhibit A.<sup>1</sup> Plaintiffs' proposed protocol (a) requires coordination between the Class Cases (in S.D.N.Y. and N.D. Cal.), and separately among the News Cases,<sup>2</sup> but *not* omnibus coordination among *all* the cases; and (b) sets sensible parameters to ensure depositions are efficient but meaningful. Defendants' proposal to have all cases coordinated sets unreasonable constraints on all Plaintiffs' ability to develop their case. *See Ex. B.* For the following reasons, Class Plaintiffs respectfully request that the Court enter their proposed deposition protocol.

***Deposition coordination across all open copyright cases against OpenAI is impractical.***

Defendants seek coordination between the Class and News Cases—literally all open copyright cases pending against OpenAI—based on the purported similarity among the cases. This is inefficient, will prejudice Plaintiffs, and ensure that the parties spend far more time on dealing with coordination issues than time spent actually deposing witnesses and moving these cases forward.

OpenAI has been sued a lot for infringement of intellectual property, much in the way that Apple or Samsung face patent infringement lawsuits. The answer has never been that Apple's Chief Technology Officer gets deposed once in all cases alleging patent infringement against the iPhone, nor is answer the right answer here to force all copyright plaintiffs—no matter how

<sup>1</sup> The parties have met and conferred about this relief and did not reach agreement.

<sup>2</sup> The Class Cases are the S.D.N.Y. Class Cases (*Authors Guild et al. v. OpenAI Inc. et al.*, (S.D.N.Y. No. 1:23-cv-8292), and *Alter et al. v. OpenAI Inc. et al.*, (S.D.N.Y. No. 1:23-cv-10211)), and the California Cases (*Tremblay, et al. v. OpenAI, Inc., et al.*, (N.D. Cal. No. 3:23-cv-03223); *Silverman, et al. v. OpenAI, Inc., et al.*, (N.D. Cal. No. 4:23-cv-03416); *Chabon, et al. v. OpenAI, Inc., et al.*, (N.D. Cal. No. 4:23-cv-04625)). The News Cases are *The New York Times Co. v. Microsoft Corp., et al.*, (S.D.N.Y. No. 1:23-cv-11195) and *Daily News, LP, et al. v. Microsoft Corp., et al.*, (S.D.N.Y. No. 1:24-cv-3285).

October 15, 2024

Page 2

different the case—to share time and coordinate deposition schedules across all open cases.

The reality is that the Class Cases, focused on books, are sufficiently different from the News Cases, focused on news articles, that coordination across those two groups of cases simply does not make sense. For one thing, the training data at issue in the Class Cases is different from the News Cases. OpenAI obtained books datasets from pirate websites. By contrast, The Times alleges that OpenAI “scraped” articles from its website. *N.Y. Times*, Dkt. 170 ¶ 160. The different training data at issue will affect not only the questions asked at depositions, but also the deponents they prioritize. The books datasets were compiled and created by particular individuals on particular teams through distinct processes. Author Plaintiffs will prioritize these deponents and questions about this particular work; News Plaintiffs may not. Those differences undercut the value of forced coordination between the two sets of cases.

In fact, Defendants agree that discovery between the Class Cases and News Cases will diverge in key ways. For instance, OpenAI admitted a few months back that discovery in the *New York Times* action would “differ substantially” from discovery in the Class Cases. *N.Y. Times*, 23-CV-11195, Dkt. 72 ¶ 42 (S.D.N.Y. (Mar. 8, 2024)). That is why OpenAI has refused to cross-produce documents produced in the *NY Times* Action in the Class Cases. See, e.g., Ex. C at 3 (refusing to produce documents produced in the *NY Times* action “because of differences between the two cases.”).

Lumping all open copyright cases against OpenAI into a single deposition protocol will also be administratively unworkable. It will force the parties and attorneys across eight different consolidated cases to identify which depositions to take and when, then force attorneys at over ten different law firms to identify times that work for both the parties and the Court. This level of coordination makes identifying a *date* for a single deposition a monumental task. But this level of coordination is not necessary. A far more logical solution is to treat the Class Cases and News Cases separately, a solution that still meaningfully minimizes the burden on OpenAI and Microsoft witnesses, while avoiding the administrative difficulties of Defendants’ kitchen-sink deposition protocol.

In cases like this one, courts favor informal coordination over formal ordered coordination. Here, that would mean that the Class Plaintiffs’ and News Plaintiffs’ make good faith efforts to schedule depositions of the same witness on back-to-back days between the two sets of cases, when feasible and possible. *Steuben Foods, Inc. v. Oystar Group*, 13-CV-1118, 2015 WL 9275748 at \*1-2 (W.D.N.Y. Dec. 21, 2015); *Or. Laborers Employers Pension Tr. Fund v. Maxar Techs.*, 19-CV-124, 2022 WL 684168 at \*3 (D. Colo. Mar. 8, 2022) (“While the Court expects counsel to . . . use their best efforts to schedule depositions on back-to-back days, ordering the same when scheduling depositions involves coordinating numerous attorney and witness schedules across multiple states, would not be practical.”); *Barrett v. Apple Inc.*, 20-CV-4812 Dkt 87, at \*3 (N.D. Cal. Feb. 25, 2022) (encouraging but not requiring parties to coordinate scheduling a witness’s multiple depositions). Plaintiffs are mindful of Defendants’ concerns about its witnesses being deposed multiple times, but multiple depositions are not uncommon for corporations sued by multiple parties asserting disparate intellectual property claims about the same or similar technology. And any burden of being deposed twice does not outweigh the prejudice to Plaintiffs. See *KGK Jewelry LLC v. ESDNetwork*, No. 11-CV-9236, 2014 WL 7333291, at \*2 (S.D.N.Y. Dec. 24, 2014) (“[T]he benefits of efficiency can never be purchased at the cost of fairness.”) (quoting

October 15, 2024

Page 3

*Malcolm v. Nat'l Gypsum Co.*, 995 F.2d 346, 350 (2d Cir. 1993)). News and Class Plaintiffs should not be required to coordinate, or at a minimum, should be allowed to take several depositions out of order as “safety valves.”

***A 280-hour cap on depositions taken by Class Plaintiffs is proportional to the needs of the case.*** These are complicated cases involving novel technologies in a rapidly evolving industry. To understand those technologies, Plaintiffs will seek testimony about how the technologies were developed and how they work, obtained through depositions. Even a single one of these cases would therefore demand an unusual number of depositions. Instead, between the Class Cases and News Cases, there are five groups of plaintiffs and two groups of defendants. A 280-hour cap is well within the range of reasonableness for two sets of coordinated, complex class actions like the Class Cases here. *E.g., In re Parmalat Sec. Litig.*, No. 04-CV-30 Dkt. 414 (S.D.N.Y. Mar. 6, 2006) (900 hours); *In re Pfizer Sec. Litig.*, No. 04-CV-9866 Dkt. 223 (S.D.N.Y. Aug. 20, 2010) (600 hours); *In re Facebook, Inc. Consumer Privacy User Profile Litig.*, 3:18-MD-2843, Dkt. 1139 (N.D. Cal. Jun. 21, 2023) (306 hours); *Wausau Bus. Ins. Co. v. Turner Const. Co.*, 204 F.R.D. 248, 250 (S.D.N.Y. 2001) (“hundreds of hours”).

Plaintiffs’ proposed hours cap is also reasonable in light of the number of relevant witnesses. Defendants have together agreed to produce documents from around 40 custodians, 16 of which are identified in Rule 26(a) disclosures. Author Plaintiffs and Defendants have subpoenaed an additional 20 third parties, on top of subpoenas served in the N.D. Cal. Cases. If just a fraction of those parties are deposed within the default seven-hour limit, depositions will exceed 280 hours. Moreover, Class Plaintiffs’ proposed cap also includes time for depositions taken under Rule 30(b)(6), including a custodial and substantive Rule 30(b)(6) deposition (i.e., two depositions each) from the OpenAI Defendants and Microsoft. Plaintiffs’ request for 280 hours is reasonable, and reflects our intention to litigate the case as efficiently as possible.

***The Court should reject Defendants’ proposed per-witness limits.*** For many of the same reasons, 8 hours per witness is unworkable. Several of Defendants’ deponents have knowledge that stretches across all aspects of the case, from obtaining copyrighted material, to the intricacies of OpenAI’s LLMs, the relationship with Microsoft, and the benefits OpenAI has reaped from its infringement. Such individuals will likely require more than eight hours of testimony, especially to multiple groups of plaintiffs. Notably, many of these people are the same individuals whose testimony Defendants seek to restrict even more: OpenAI’s C-Suite executives. Limiting those witnesses’ depositions would severely hamstring Plaintiffs’ case. It would moreover be premature: if Defendants claim deposing a particular witness is unduly burdensome, they should substantiate that claim with respect to each witness by seeking a protective order at the appropriate time, rather than seeking blanket protection through a deposition coordination.

In any event, restrictive per-witness caps are unnecessary here. The “tight time limits . . . herein provide adequate incentive for all parties to conduct depositions as expeditiously as possible.” *Parmalat*, Dkt. 414; *see Pfizer* Dkt. 223. Plaintiffs’ proposal—10 hours among the Class Cases—strikes a fair middle ground; if the Court orders Author Plaintiffs to coordinate with News Plaintiffs, that number should be expanded further. *Pierre v. City of New York*, 2022 WL 2384150, at \*1 (S.D.N.Y. July 1, 2022) (granting additional deposition for a witness because, in “multi-party cases, the need for each party to examine the witness may warrant additional time”).

October 15, 2024

Page 4

***Defendants' proposed limitations on Rule 30(b)(6) depositions are improper and should be rejected.*** Defendants ask that Plaintiffs in the Class and News Cases take a complete Rule 30(b)(6) deposition of the Defendants in 21 hours. That is not possible. The details of Defendants' proposal only magnify those issues. Defendants allow themselves almost unlimited leeway to limit Rule 30(b)(6) time. To take a few examples: Because they offer no minimum for 30(b)(6) time, Defendants could designate a single 30(b)(6) witness on all topics across all cases, requiring under their proposal that all Plaintiffs to complete 30(b)(6) questioning in 12 hours. Worse, their unified hours cap for personal and representative depositions would let Defendants "curtail 30(b)(6) examinations by designating as a 30(b)(6) witness a person who previously testified for [8] hours as an individual and has only [4] hour[s] left on his or her presumptive [12]-hour clock," restricting 30(b)(6) time to a total of 4 hours across all cases. *Sabre v. First Dominion Capital*, 01-CV-2145, 2001 WL 1590544 at \*1 (S.D.N.Y. Dec. 10, 2001). Conversely, "to limit the testimony of an individual," Defendants could "designat[e] the individual as a 30(b)(6) witness; under [Defendants' proposal,] every minute spent conducting the 30(b)(6) deposition would be deducted from the time available to probe the witness's individual knowledge." *Id.* These attempts to shield themselves from questioning are directly contrary to the Federal Rules, *id.*; Plaintiffs' proposal addresses them both.

***Plaintiffs should be permitted to take preliminary custodial Rule 30(b)(6) depositions of Defendants.*** This will significantly streamline discovery. It was discussed at the parties' most recent conference, *see* Sep. 12 Hearing Tr. 41, 51, 82-83, and even ordered by Judge Alsup in a similar LLM copyright case in the N.D. Cal. Ex. D at 10-11. Early custodial 30(b)(6) depositions are a common practice in large cases. *See, e.g., Pfizer*, Dkt. 223 (authorizing the parties to take "30(b)(6) depositions concerning non-case specific issues such as corporate structure, corporate filing systems, etc." early); *In re Zoloft*, 2:12-MD-2342, Dkt. 287 (E.D. Pa. Nov. 16, 2012) (ordering "initial 30(b)(6) depositions on ESI [and] corporate organization/structure"); *Durst v. Zimmer Biomet Holdings, Inc.*, 19-CV-790, Dkt. 6 (S.D.N.Y. Jan. 20, 2019) (ordering an early 30(b)(6) deposition "on the topic of corporate organization and ESI"). And it is particularly critical here, as Defendants have consistently failed to answer basic questions about repositories, *e.g.*, Ex. E, or produce foundational documents evidencing corporate structure. *E.g.*, Ex. F. An early 30(b)(6) deposition will provide needed clarity and limit the disputes the parties bring to the Court.

***Defendants may not tell Plaintiffs how to spend their hours.*** Defendants seek to limit the number of hours Plaintiffs may depose each Defendant—that is, to limit how much time can be spent examining OpenAI witnesses versus Microsoft witnesses. There is no reason why Defendants should have a say on how Plaintiffs use their time under the deposition cap. Such limits would "hamstring Plaintiffs' ability to prove their case . . . . Plaintiffs' proposal, by contrast, preserves the ability of both sides to make deposition decisions based on their specific needs to prove their claims or defenses." *United States v. Am. Airlines Grp.*, 1:21-CV-11558, Dkt. 65 (D. Mass. Nov. 19, 2021) and Dkt. 75 (D. Mass. Nov. 29, 2021) (rejecting defendants' attempt to "dictate how Plaintiffs must allocate their depositions").

***250 hours for Defendants to take depositions is sufficient here.*** Defendants argue that more hours are warranted for depositions of Plaintiffs than of Defendants. Not so. This case will rise or fall on Defendants' conduct—whether OpenAI's mass reproduction and exploitation of authors' works constituted copyright infringement, whether OpenAI conduct constitutes fair use,

October 15, 2024

Page 5

whether authors' works were in OpenAI training dataset, and whether Microsoft knew of the contents of OpenAI's training dataset. While Plaintiffs are the authors and owners of books allegedly used to train OpenAI's models, they were not involved in the creation of OpenAI's large language models or other issues principally in dispute in this case. Given their limited direct involvement in the central issues in the case, class plaintiff depositions regularly take only 3 or 4 hours each. *Carlin v. Daily Am., Inc.*, 09-CV-430, Dkt. 542 (E.D. Cal. Jan. 29, 2018) (limiting class plaintiff depositions to 4 hours); *Onate v. AHRC Health Care, Inc.*, 20-CV-8292, 2023 WL 6973734 at \*3 (S.D.N.Y. Oct. 23, 2023) (3 hours); *Severino v. Avondale Care Grp.*, 1:21-CV-10720, Dkt. 259 at \*2 (S.D.N.Y. May 16, 2023) (4 hours). Even if Defendants depose each of those individuals for 4 hours, at 250 hours, they will have over 50 hours left over for further questioning or additional deponents. 250 hours is more than enough for Defendants to build out their defenses.

***Defendants must cross-produce documents across coordinated cases.*** To the extent the Court requires coordination, whether with the N.D. Cal. case or the News Cases, Defendants must cross-produce documents across the cases in which depositions are coordinated. Defendants contend that cross-production is "outside the scope of this deposition protocol." Ex. B. To the contrary, those issues go hand-in-hand. First, coordination assumes similarities between the cases, and therefore overlapping documents. And second, Plaintiffs cannot efficiently coordinate depositions unless they are making decisions based on the same information. That is why cross-production is standard where parties are coordinating discovery. See *Brown v. Google*, 20-CV-3664, Dkt. 243 (N.D. Cal. Aug. 12, 2021) (requiring defendant to "produce across both Matters all documents of Overlapping Custodians that are produced in either action" to facilitate deposition coordination). For that reason, if the Court requires coordination with the N.D. Cal. Cases, OpenAI should cross-produce "all documents and written discovery," with two exceptions: documents that relate only to the California Unfair Competition Law, and documents specific to Microsoft's liability. Ex. A. By carving out those documents, which are relevant only to the N.D. Cal. and S.D.N.Y. cases, respectively, Author Plaintiffs are seeking production only of documents relevant to their case, consistent with Judge Illman's decision in the N.D. Cal. action. *Tremblay*, 3:23-CV-3223, Dkt. 140 (N.D. Cal. May 14, 2024). By the same token, if the Court orders coordination between the News Cases and this one, OpenAI and Microsoft must cross-produce any documents that relate to both cases.

For the foregoing reasons, the Court should enter Plaintiffs' proposed deposition coordination protocol.

Sincerely,

LIEFF CABRASER HEIMANN SUSMAN GODFREY LLP COWAN, DEBAETS,  
& BERNSTEINS LLP ABRAHAMS & SHEPPARD  
LLP

/s/ Rachel Geman  
Rachel Geman

/s/ Rohit Nath  
Rohit Nath

/s/ Scott J. Sholder  
Scott J. Sholder